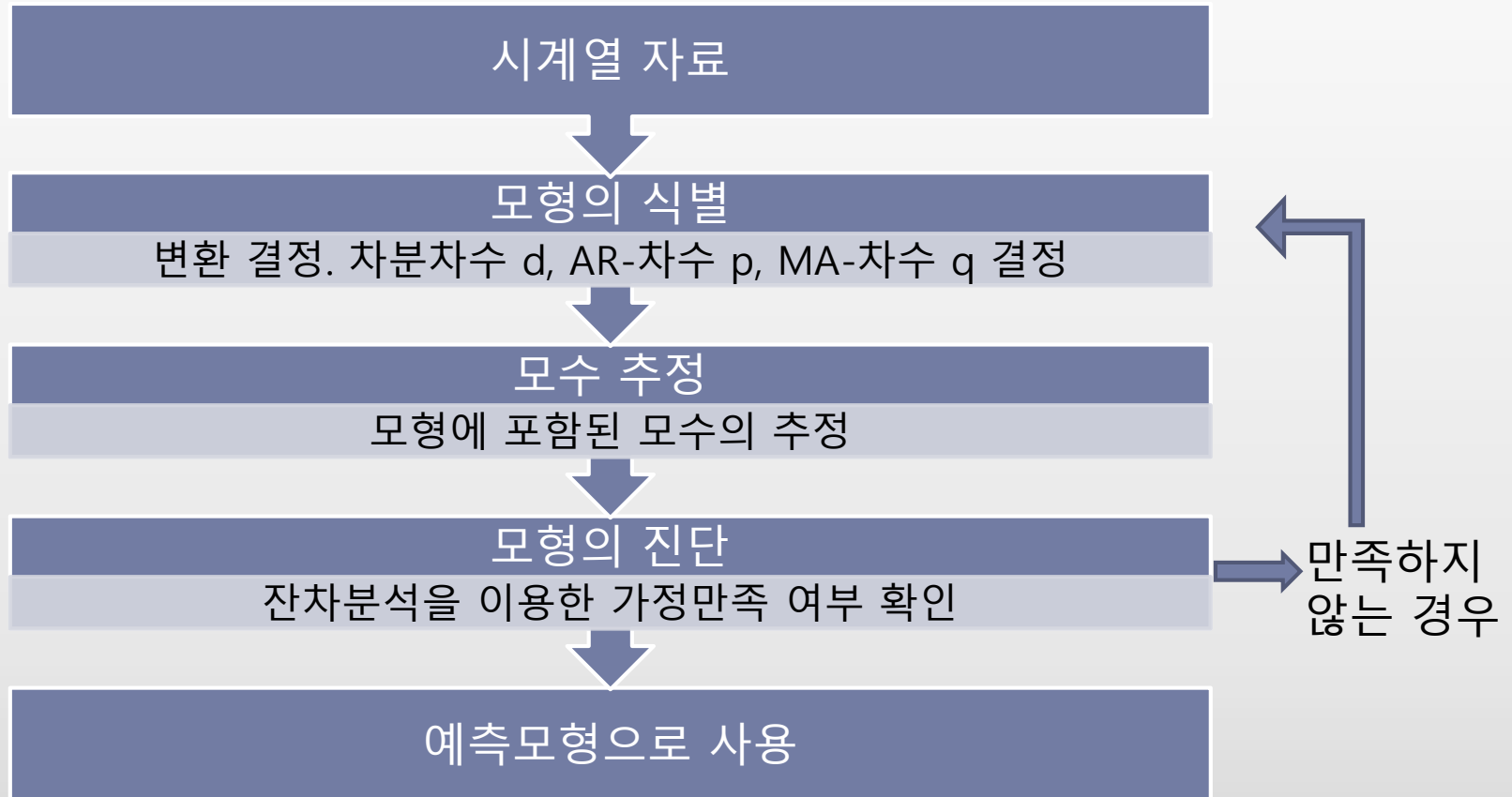


## 8장. ARIMA 모형의 적합

모형적합 절차. 모형의 식별. 모수의 추정. 모형의 진단

## 8.1 모형의 적합절차



## 8.2 모형의 식별

---

- ARIMA(p,d,q) 모형에서 차분 차수 d, AR 차수 p, MA 차수 q 결정
- 모수 간결의 원칙: 가능한  $p \leq 2$  ,  $q \leq 2$  유지

## 모형 식별 1 단계: 시계열 그림과 SACF 작성

---

- 시계열 정상화 단계
- 동일분산 확인: 필요 시 분산안정화 변환 실시
- 추세 혹은 계절 성분 확인: 필요 시 차분 실시
  - 최적 차분 차수의 결정: ACF의 형태 혹은 단위근 검정 결과로 판단
  - 대부분의 경우 2차 이상의 차분은 필요 없음
  - 과대 차분 여부 확인

## 모형 식별 2 단계: ARMA 모형의 차수 p와 q 결정

---

- SACF 및 SPACF를 이용하여 차수 p와 q를 결정
  - SACF를 절단으로 인식: MA 모형. 절단 시점이 차수 q
  - SPACF를 절단으로 인식: AR 모형. 절단 시점이 차수 p
  - SACF와 SPACF 모두 감소로 인식: ARMA 모형.
    - ARMA(p,q),  $p \leq 2$ ,  $q \leq 2$ 의 모형 중에서 AIC 혹은 SBC(BIC) 값이 최소 모형 선택

$$AIC = -2 \times \log L + 2(p + q)$$

$$SBC = -2 \times \log L + \log(n) \times (p + q)$$

$\log L$ : log likelihood

## 모형 식별 3 단계: 절편 $\delta$ 모형에 포함 여부 결정

---

- 차분을 실시하지 않은 경우:
  - 함수 `arima()`에서 옵션 `include.mean=TRUE`(디폴트)로  $H_0: \delta=0$ 을 검정
  - 귀무가설을 기각할 수 없게 되면 `include.mean=FALSE`로 절편 제거
- 차분을 실시한 경우:
  - 함수 `arima()`에서 결정할 수 없음
  - 대안 1: 차분된 자료를 대상으로 모평균이 0인지 여부를 직접 검정

$$T = \frac{\bar{W}}{S_w / \sqrt{n}}, \quad W_t = (1-B)^d Z_t \quad \text{교재 281쪽 참조}$$

- 대안 2: `forecast::Arima()` 이용

## 8.3 모수 추정

---

- 모수 추정 방법
  - 조건부 최소제곱 추정법(minimizing conditional sum-of-squares)
  - 비조건부 최소제곱 추정법(minimizing unconditional SS)
  - 최대가능도 추정법(maximum likelihood estimation)
- 어떤 방법을 선택?
  - 이론적으로 각 방법마다 장단점이 있음
  - 대부분의 경우 추정 결과에는 큰 차이가 없음
  - 함수 `arima()`의 디폴트 방법("CSS-ML") 사용

## 8.4 모형 진단

---

- 모형 식별과 모수 추정을 통해 얻어진 잠정모형의 타당성 여부를 확인하는 단계
- 모형 진단
  - 잔차분석
  - 과대적합
- 잔차분석: 오차항이 정규분포 백색잡음과정을 따르는지 여부를 확인. 회귀분석의 절차와 유사.
- 과대적합: 잠정모형에 모수를 추가한 모형의 유의성을 확인하는 분석



## 8.4.1 잔차분석

---

- 정상 ARMA(p,q) 모형:  $\phi(B)(Z_t - \mu) = \theta(B)\varepsilon_t$ 
  - 오차항  $\varepsilon_1, \varepsilon_2, \dots$  는 서로 독립이고 평균이 0, 분산이  $\sigma^2$ 인 정규분포
- 가정 만족 여부 확인 방법
  - 정규성 확인: 정규 Q-Q plot
  - 평균 0, 동일 분산 확인: 잔차의 시계열 그림
  - 서로 독립 확인:
    - 1) 잔차의 SACF(RSCAF) 및 SPACF(RSPACF)
    - 2) 포트맨토 검정: 귀무가설은 잘 정의되어 있으나 대립가설은 상대적으로 느슨하게 정의된 가설에 대한 검정

## 시계열 분석에서의 포트맨토 검정

---

- 귀무가설:  $H_0: \rho_1(e) = \rho_2(e) = \dots = \rho_K(e) = 0$ 
  - 시차 K까지 잔차들 사이에는 자기상관이 없다.
  - 대립가설: 귀무가설은 사실이 아니다. 즉, 독립이 아니다.  
구체적인 상관관계의 구조를 명시하고 있지 않음.

- ARMA(p,q) 모형에 대한 검정

- Box-Pierce Test

$$Q^* = n \sum_{k=1}^K \hat{\rho}_k^2(e) \sim \chi^2(K-p-q) \quad \text{Under } H_0$$

- Ljung-Box Test: 소규모 데이터에 더 적합한 검정

$$Q = n(n+2) \sum_{k=1}^K \hat{\rho}_k^2(e) / (n-k) \sim \chi^2(K-p-q) \quad \text{Under } H_0$$

## 8.4.2 과대적합

---

- 잠정모형에 모수를 추가하여 더 많은 개수의 모수를 갖는 모형을 적합시키는 것
- 과대적합 실시 이유:
  - 잠정모형이 적절한 모형이라면 추가된 모수들은 유의하지 않게 나올 것이기 때문
- 모수 추가 방법: 잠정모형에 하나의 AR항과 MA항을 각각 추가하여 두 개의 과대적합모형 설정
  - 주의 사항: AR항과 MA항을 동시에 추가해서는 안 됨

## 과대적합의 예

---

- 만일 AR(1) 모형이 잠정모형으로 선택되었다면 과대적합 모형으로 AR(2)와 ARMA(1,1) 모형을 각각 적합
- 잠정모형의 교체
  - 1) 추가된 모수, 즉 AR(2)의  $\phi_2$  혹은 ARMA(1,1)의  $\theta$ 가 유의적
  - 2) 기존의 모수, 즉 AR(2)의  $\phi_1$  혹은 ARMA(1,1)의  $\phi$ 에 대한 추정 결과에 큰 변화
  - 3) 과대적합 모형에서의 잔차 분산이 더 작아질 때

→ 추가된 모수가 설명력이 있다고 판단  
과대적합 모형을 잠정모형으로 지정

# ARIMA 모형의 적합에 필요한 R 함수

---

- 최적 차분 차수 결정
  - 함수 `forecast :: ndiffs( )` : 비계절형 차분의 경우
- ARIMA(p,d,q) 모형 적합
  - 최적 d, p, q 차수를 모르는 경우: 함수 `forecast :: auto.arima( )`
  - 최적 차수를 아는 경우: 함수 `arima( )`
- 모형 진단
  - 함수 `tsdiag( )`: 잔차 시계열 그림, 잔차의 ACF, 포트맨토 검정 결과(p값)

## 1) 함수 ndiffs( )의 사용법

---

- 단위근 검정으로 정상성을 확보할 수 있는 차분 차수 결정
- 패키지 forecast의 함수

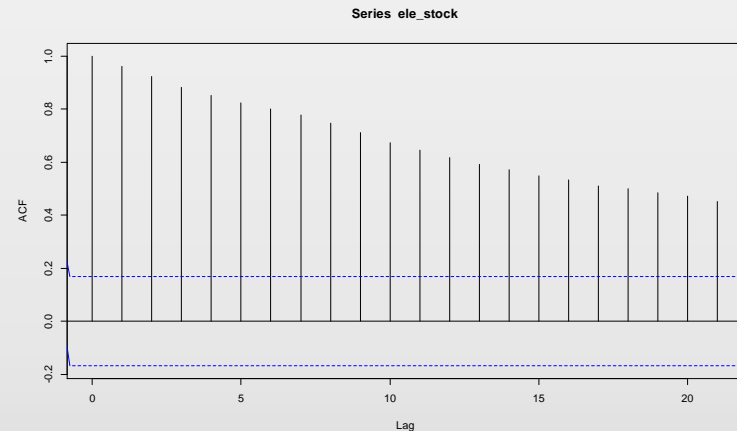
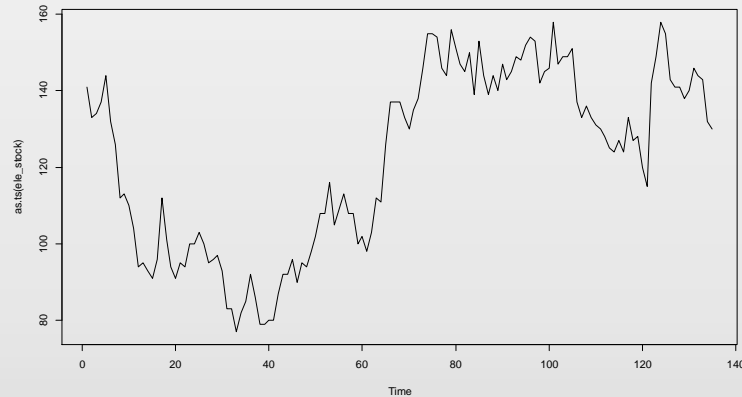
`ndiffs(x, alpha=0.05, test=c("kpss", "adf", "pp"))`

- x: 시계열 자료
- alpha: 유의수준
- test="kpss": KPSS 검정. 귀무가설- 정상시계열, 대립가설- 비정상시계열
- test="adf" : ADF 검정. 귀무가설-비정상 시계열, 대립가설- 정상시계열
- test="pp" : PP 검정. 가설은 ADF 검정과 동일

## 예제: 교재 316쪽 예8-8

### - 모 전자회사의 주별 주가지수

```
> ele_stock <- scan("D:/Data/elecstock.txt")  
Read 135 items  
> plot(as.ts(ele_stock))  
> acf(ele_stock)
```

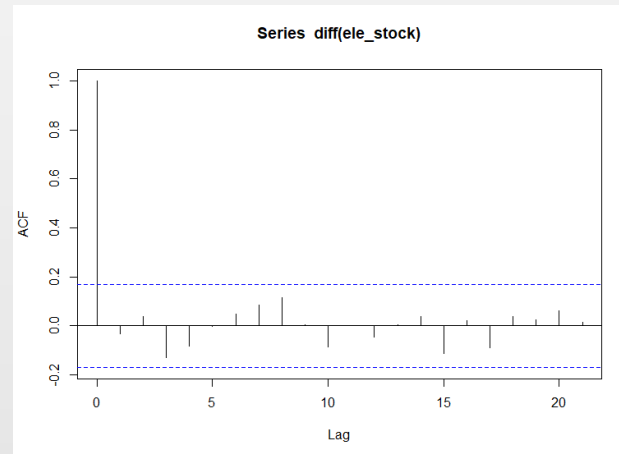
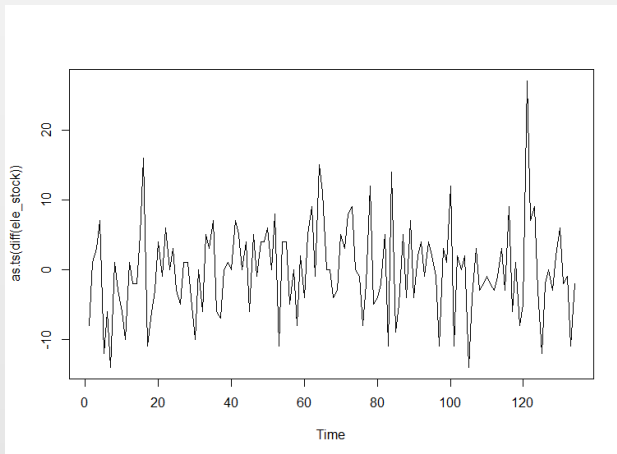


비정상 시계열

- 함수 `ndiffs()`에 의한 최적 차분 결정

```
> library(forecast)
> ndiffs(ele_stock)
[1] 1
```

## 1차 차분된 시계열 자료의 시계열 그림과 ACF



1차 차분으로 정상성 확보



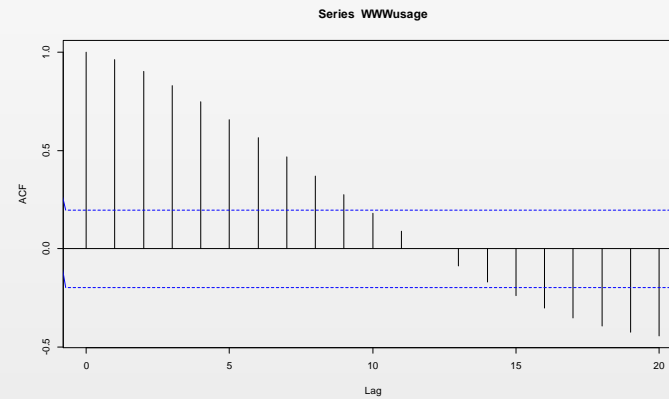
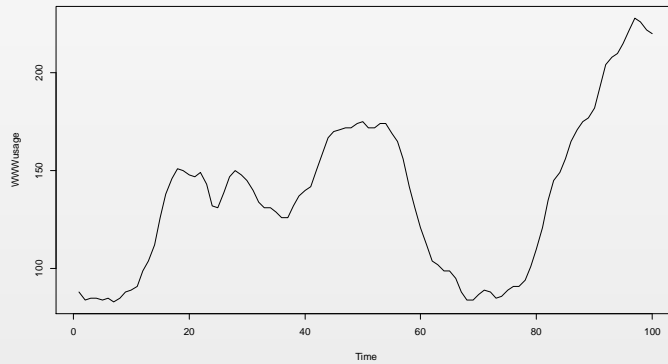
## 2) 함수 `auto.arima()`의 사용법

---

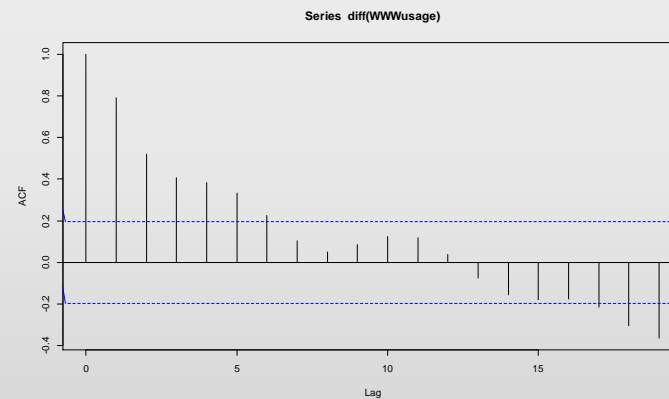
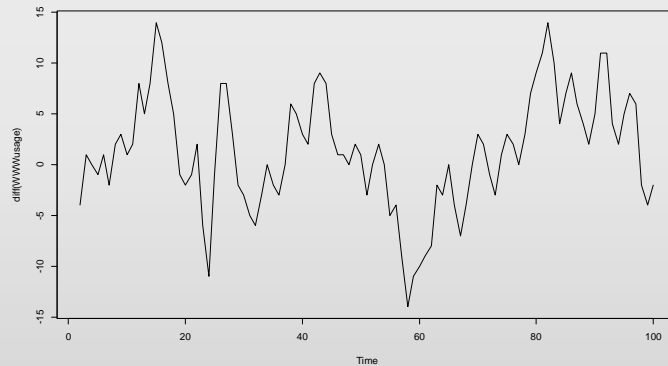
- AIC 혹은 BIC를 기준으로 최적 ARIMA(p,d,q) 모형 적합
- 패키지 `forecast`의 함수
- 사용법: `auto.arima(x)`, x: 시계열 자료
- 선택한 모형이 최종모형은 아님
- 모형 인식의 가이드 라인 제시

# 예제: WWWusage

## - 분당 인터넷 접속자 수



## - 1차 차분된 자료



- 함수 `auto.arima()`로 모형 인식

```
> library(forecast)

> auto.arima(wwwusage)
Series: wwwusage
ARIMA(1,1,1)

Coefficients:
           ar1      ma1
       0.6504  0.5256
s.e.  0.0842  0.0896

sigma^2 estimated as 9.995:  log likelihood=-254.15
AIC=514.3    AICc=514.55    BIC=522.08
```

### 3) 함수 `arima( )`의 사용법

---

- `ARIMA(p,d,q)` 모형의 적합
- `arima( x, order=c(0, 0, 0) , include.mean=TRUE , fixed=NULL )`
  - `x`: 시계열 자료
  - `order=c(p, d, q)`의 순서로 차수 지정
  - `include.mean`: 절편  $\delta$  포함 여부.  $d=0$ 의 자료에 대해서는 `TRUE`가 디폴트.  $d \geq 1$ 의 자료에 대해서는 적용 불가능.
  - `fixed`: 비유의적인 모수를 모형에서 제거할 때 사용되는 옵션. 예제와 함께 설명할 예정.

## 4) 함수 `tsdiag()`의 사용법

---

- 모형 진단에 유용하게 사용되는 함수
- `tsdiag( object , gof.lag , ... )`
  - `object`: 적합된 시계열 모형 객체. 함수 `arima()` 등으로 생성된 객체
  - `gof.lag`: 포트맨토 검정에서 사용될 최대 시차
- 결과: 그래프 출력
  - 잔차 시계열 그림
  - 잔차의 ACF
  - 옵션 `gof.lag`에서 지정한 시차  $K$ 까지의 다음의 가설에 대한 Ljung-Box 검정의  $p$ 값

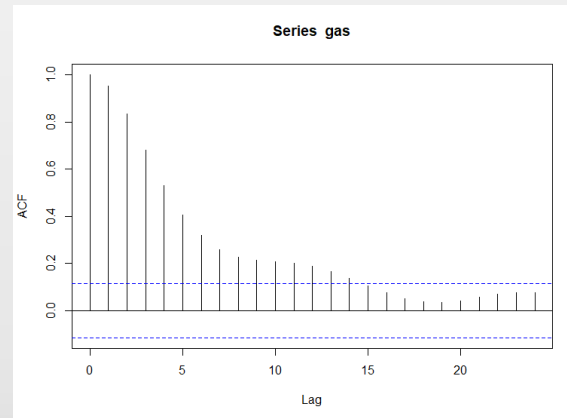
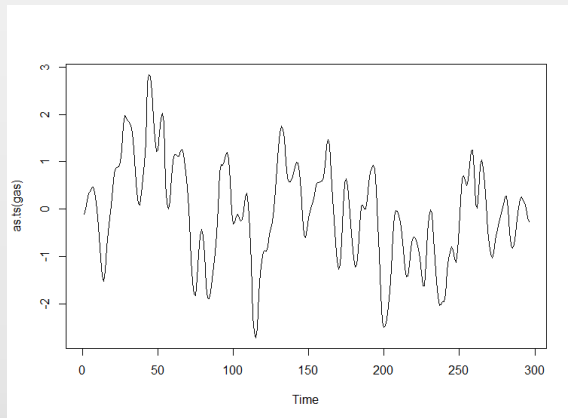
$$H_0: \rho_1(e) = \rho_2(e) = \dots = \rho_K(e) = 0$$

## 8.5 모형 적합 예제

- 예 8-6 (교재 308쪽): 데이터 파일 gas.txt

```
> gas <- scan("D:/Data/gas.txt")  
Read 296 items
```

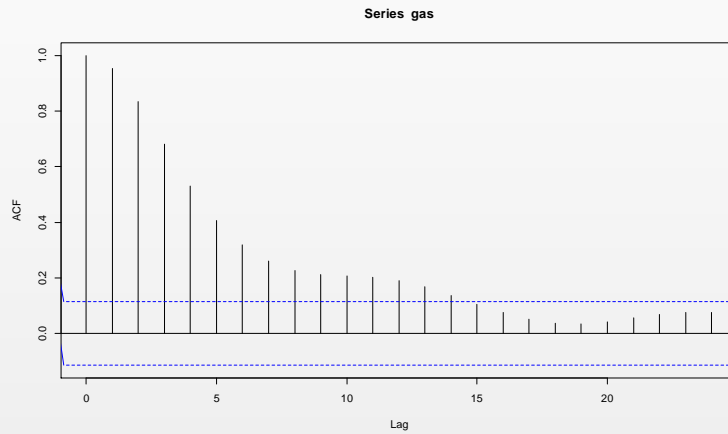
- 정상성 만족 여부 확인



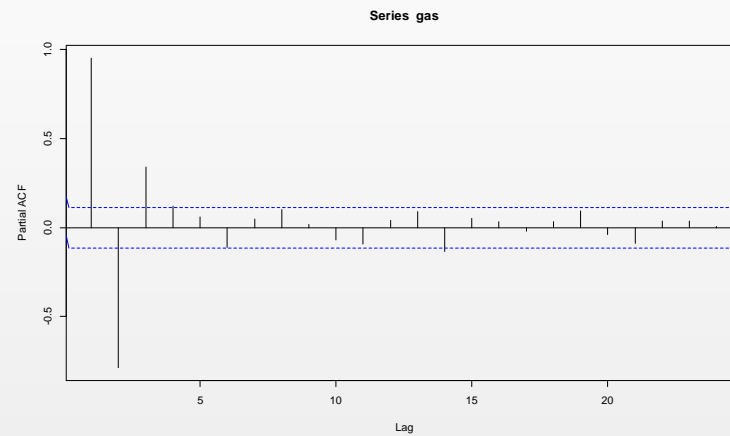
```
> library(forecast)  
> ndiffs(gas)  
[1] 1
```

차분 차수  $d=0$ 과  $d=1$ 의 경우를  
모두 확인할 필요가 있음

◆ 차분 차수  $d=0$ 의 경우: 모형 인식



ACF 감소



PACF 절단  
또는  
PACF 감소

모형 인식

- AR(3)
- ARMA(1,1), ARMA(1,2)
- ARMA(2,1), ARMA(2,2)

## 1) AR(3) 모형 적합

```
> fit <- arima(gas, order=c(3,0,0))  
> fit
```

Coefficients:

	ar1	ar2	ar3	intercept
	1.9691	-1.3651	0.3394	-0.0606
s.e.	0.0544	0.0985	0.0543	0.1898

sigma^2 estimated as 0.0353: log likelihood = 72.57, aic = -135.14

### - 절편 유의성 확인

```
> confint(fit)
```

	2.5 %	97.5 %
ar1	1.8624733	2.0756582
ar2	-1.5582740	-1.1720122
ar3	0.2329228	0.4458861
intercept	-0.4326438	0.3113574



- 절편을 제외한 모형 적합

```
> fit <- arima(gas,order=c(3,0,0),include.mean=FALSE)
> fit
```

call:

```
arima(x = gas, order = c(3, 0, 0), include.mean = FALSE)
```

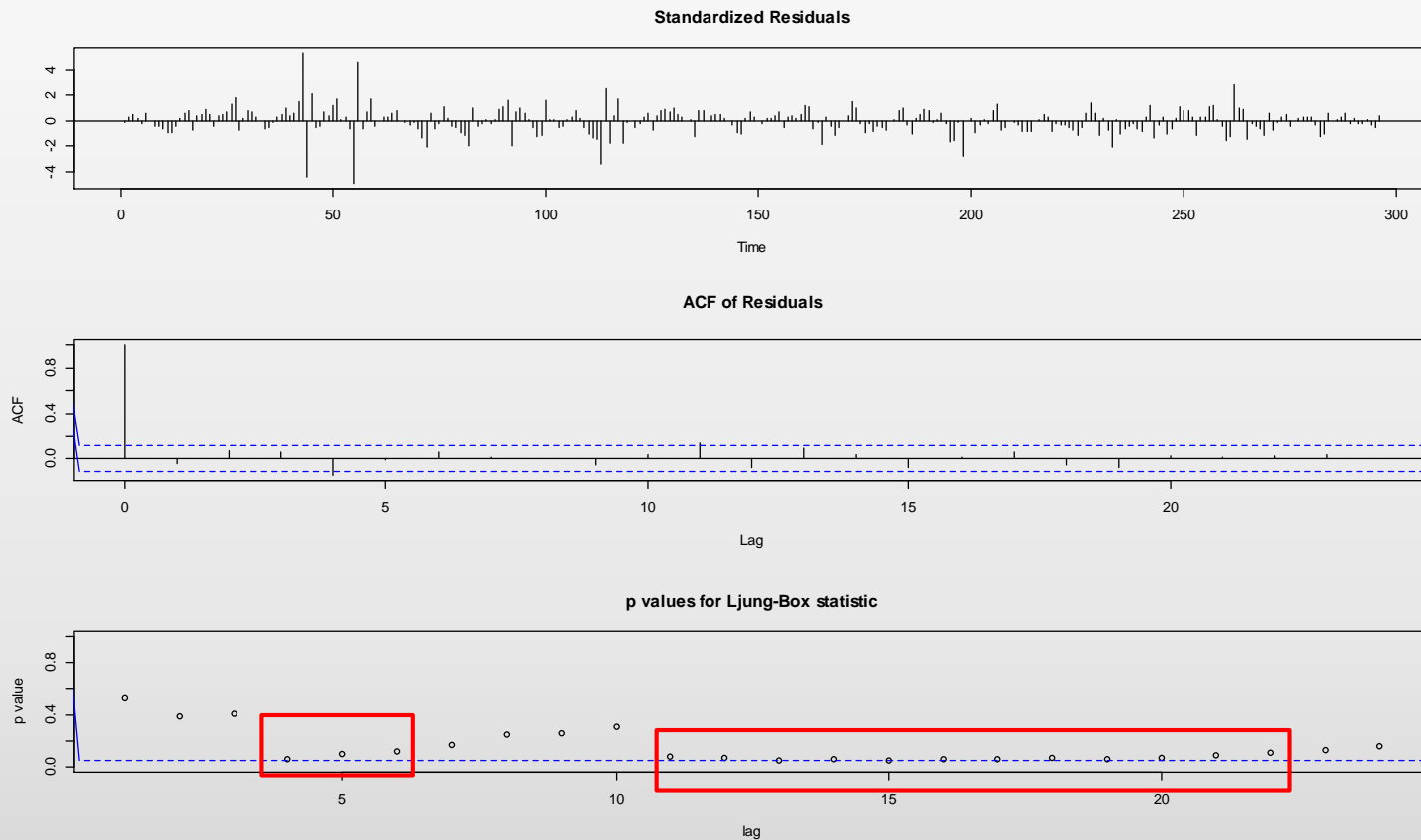
Coefficients:

	ar1	ar2	ar3
	1.9696	-1.3659	0.3399
s.e.	0.0544	0.0985	0.0543

sigma^2 estimated as 0.03531: log likelihood = 72.52, aic = -137.04

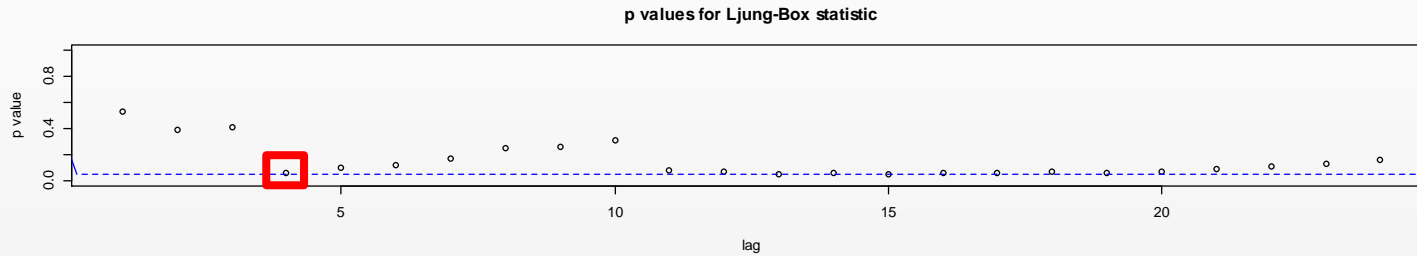
- AR(3) 모형에 대한 모형 진단

```
> fit <- arima(gas,order=c(3,0,0),include.mean=FALSE)
> tsdiag(fit,gof.lag=24)
```



부적절한  
모형

- 함수 `tsdiag()` 포트맨토 검정의 문제



- 상세한 검정 결과 산출: `Box.test()`

```
> Box.test(fit$resid, lag=4, type="Ljung-Box")
data:  fit$resid
X-squared = 9.1435, df = 4, p-value = 0.05761
```

$df = K - p - q$

- $K=4, p=3$
- $df=1$
- wrong df 사용

```
> Box.test(fit$resid, lag=4, type="Ljung-Box", fitdf=3)
data:  fit$resid
X-squared = 9.1435, df = 1, p-value = 0.002496
```

함수 `Box.test()`의 옵션 `fitdf` 추가:  $fitdf = p + q$

- AR(3) 모형에 대한 과대적합

```
> fit1.1 <- arima(gas,order=c(4,0,0),include.mean=FALSE)
> confint(fit1.1)
                2.5 %      97.5 %
ar1  1.814736086  2.0400206
ar2 -1.444161417 -0.9507063
ar3 -0.148252749  0.3450517
ar4  0.009420644  0.2343153

> fit1.2 <- arima(gas,order=c(3,0,1),include.mean=FALSE)
> confint(fit1.2)
                2.5 %      97.5 %
ar1  2.0242296  2.47743254
ar2 -2.2271911 -1.46281370
ar3  0.3838359  0.73953535
ma1 -0.5844018 -0.05805505
```

AR(4)

- 추가된 모수
- 유의적
- 비유의적
- 모수 포함
- 추가 분석

ARMA(3,1)

- 모든 모수
- 유의적
- 추가분석

- AR(4) 모형에서 비유의적 모수 제거

```
> fit1.1 <- arima(gas,order=c(4,0,0),fixed=c(NA,NA,0,NA),
  include.mean=FALSE)

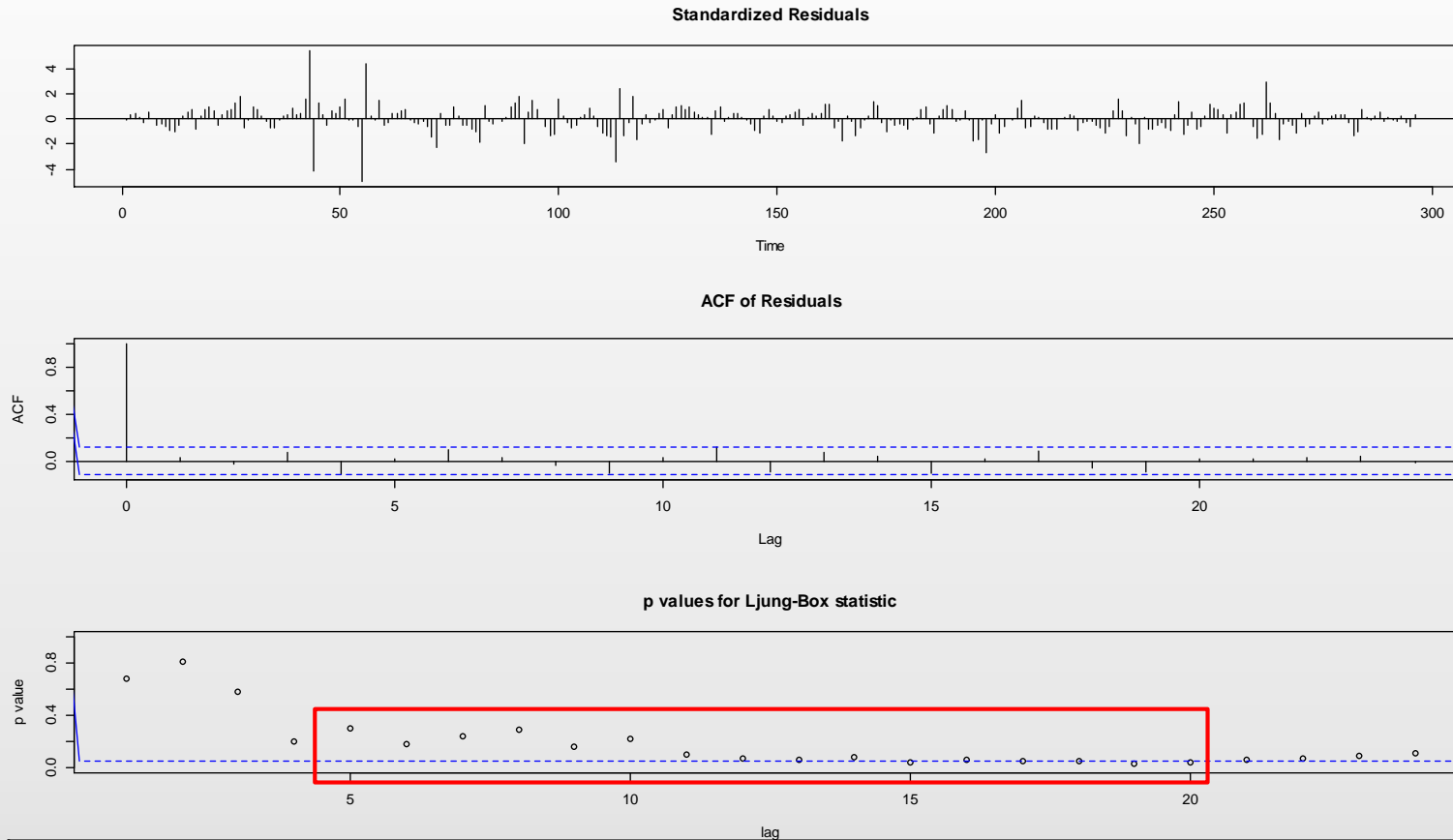
> fit1.1

Coefficients:
      ar1      ar2      ar3      ar4
    1.8990 -1.1105      0  0.1624
s.e.  0.0446  0.0592      0  0.0246

sigma^2 estimated as 0.03484:  log likelihood = 74.45,
aic = -140.9
```

- 자세한 사용법은 다시 설명할 예정
- 경고 메시지는 무시

- AR(4) 모형의 모형진단



부적합한  
모형

```
> Box.test(fit1.1$resid, lag=5, type="Ljung-Box", fitdf=4)$p.value
[1] 0.01388278
```

- AR(4) 모형의 과대적합

AR 모수

MA 모수

```
> confint(arima(gas,order=c(4,0,1),include.mean=FALSE,
               fixed=c(NA,NA,0,NA,NA)))
```

	2.5 %	97.5 %
ar1	1.71631440	2.0176740
ar2	-1.26424832	-0.8756488
ar3	NA	NA
ar4	0.08249962	0.2181195
ma1	-0.14196767	0.2474513

```
> confint(arima(gas,order=c(5,0,0),include.mean=FALSE,
               fixed=c(NA,NA,0,NA,NA)))
```

	2.5 %	97.5 %
ar1	1.78837847	1.9868923
ar2	-1.23779832	-0.9365919
ar3	NA	NA
ar4	-0.02268358	0.2782971
ar5	-0.07497349	0.1230255

AR(4)를 기반으로 하는 모형 탐색 중지

- ARMA(3,1) 모형 적합결과

```
> fit1.2 <- arima(gas,order=c(3,0,1),include.mean=FALSE)
> fit1.2

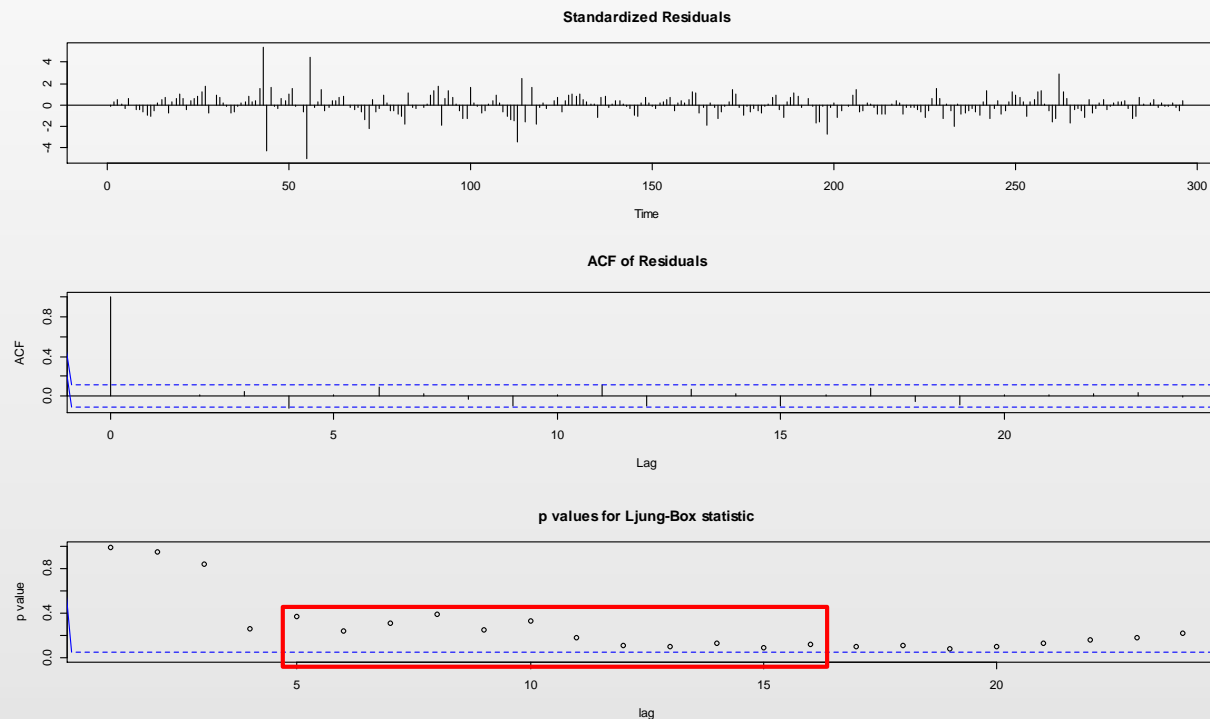
Coefficients:
            ar1      ar2      ar3      ma1
      2.2508  -1.845   0.5617  -0.3212
s.e.  0.1156   0.195   0.0907   0.1343

sigma^2 estimated as 0.03476:  log likelihood = 74.82,
aic = -139.65
```



- ARMA(3,1) 모형에 대한 모형 진단: 잔차분석

```
> tsdiag(fit1.2, gof.lag=24)
```



부적절한 모형

```
> Box.test(fit1.2$resid, lag=5, type="Ljung-Box", fitdf=4)$p.value  
[1] 0.02017975
```

- ARMA(3,1) 모형에 대한 모형 진단: 과대적합

```
> confint(arima(gas,order=c(4,0,1),include.mean=FALSE))
              2.5 %      97.5 %
ar1  1.4909061  2.7158914
ar2 -2.7661390 -0.3183116
ar3 -0.5487845  1.2182624
ar4 -0.1804910  0.3093737
ma1 -0.7868099  0.4270632

> confint(arima(gas,order=c(3,0,2),include.mean=FALSE))
              2.5 %      97.5 %
ar1  1.7860045  2.4974011
ar2 -2.2424955 -1.1390926
ar3  0.2693742  0.7417259
ma1 -0.5729986  0.1219732
ma2 -0.1547244  0.3609010
```

- 두 모형 모두 추가된 모수가 비유의적
- ARMA(3,1)을 기반으로 하는 모형 탐색 중지

→ AR(3)을 기반으로 하는 최적모형 탐색 불가능

## 2) ARMA 모형의 적합

```
> arima(gas,order=c(1,0,1),include.mean=FALSE)$aic  
[1] -25.92449  
> arima(gas,order=c(1,0,2),include.mean=FALSE)$aic  
[1] -86.0693  
> arima(gas,order=c(2,0,1),include.mean=FALSE)$aic  
[1] -127.355  
> arima(gas,order=c(2,0,2),include.mean=FALSE)$aic  
[1] -130.3612
```

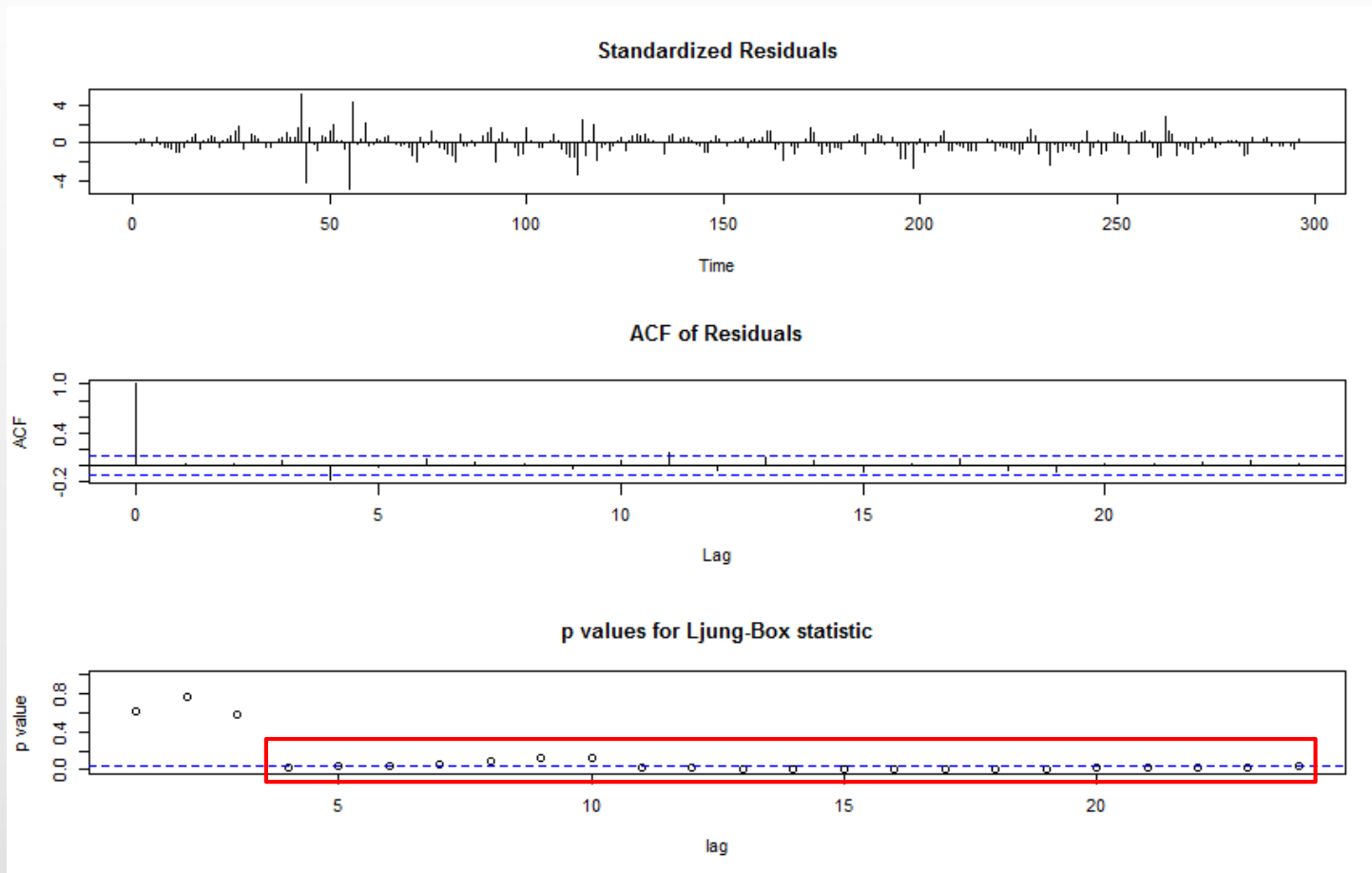
AIC 최소 모형: ARMA(2,2) 잠정모형으로 선택

- ARMA(2,2) 모형의 추정

```
> fit2 <- arima(gas, order=c(2,0,2), include.mean=FALSE)
> fit2
Coefficients:
          ar1          ar2          ma1          ma2
      1.5237   -0.6241    0.3873    0.1656
s.e.   0.0634    0.0614    0.0736    0.0716

sigma^2 estimated as 0.03588:  log likelihood = 70.18,  aic =
-130.36
```

- ARMA(2,2) 모형의 검진



부적절한 모형

- ARMA(2,2) 모형의 과대적합

```
> confint(arima(gas,order=c(3,0,2),include.mean=FALSE))
          2.5 %      97.5 %
ar1  1.7860045  2.4974011
ar2 -2.2424955 -1.1390926
ar3  0.2693742  0.7417259
ma1 -0.5729986  0.1219732
ma2 -0.1547244  0.3609010

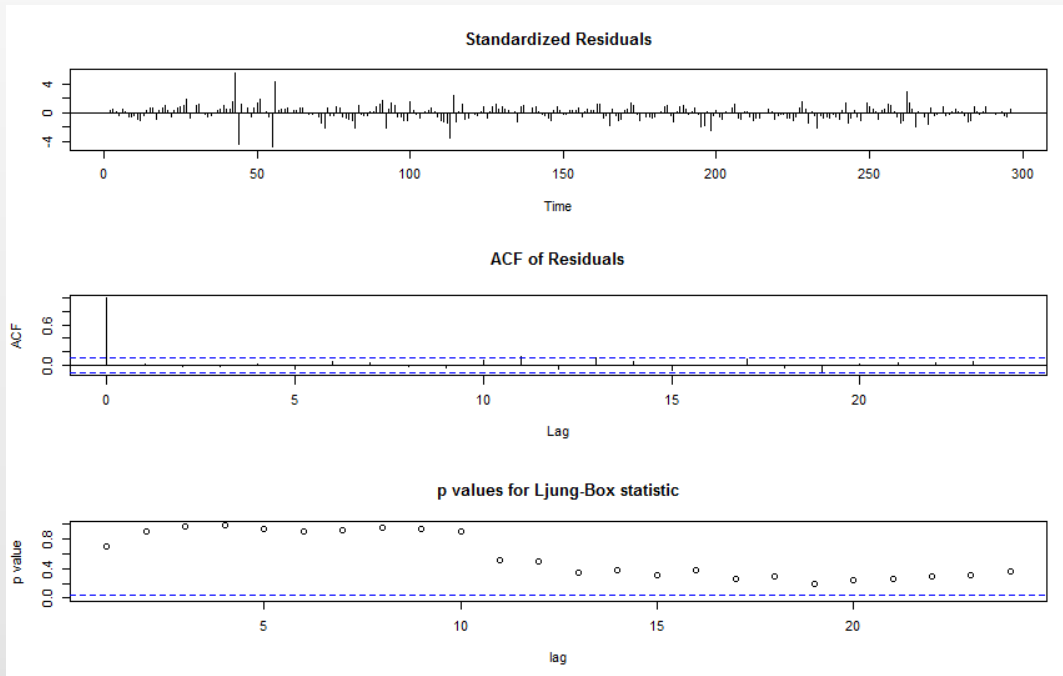
> confint(arima(gas,order=c(2,0,3),include.mean=FALSE))
          2.5 %      97.5 %
ar1  1.0950502  1.4555658
ar2 -0.5907764 -0.2516648
ma1  0.4685457  0.8045071
ma2  0.3380935  0.6827114
ma3  0.2113569  0.5115029
```

ARMA(3,2)  
의미 없는 모형

ARMA(2,3)  
모든 모수 유의적  
→ 새로운 잠정 모형

- ARMA(2,3) 모형의 검진

```
> fit2.1 <- arima(gas, order=c(2,0,3),include.mean=FALSE)
> tsdiag(fit2.1,gof.lag=24)
```



오차항 가정 만족?

```
> Box.test(fit2.1$resid,lag=19,type="Ljung-Box",fitdf=5)

data:  fit2.1$resid
X-squared = 24.093, df = 14, p-value = 0.04466
```

- ARMA(2,3) 모형의 과대적합

```
> confint(arima(gas,order=c(3,0,3),include.mean=FALSE))
              2.5 %      97.5 %
ar1  1.1508226  1.8358907
ar2 -1.3606211 -0.2950347
ar3 -0.0366154  0.4656861
ma1  0.1056627  0.7588545
ma2  0.3104800  0.6818452
ma3  0.2001320  0.5092310

> confint(arima(gas,order=c(2,0,4),include.mean=FALSE))
              2.5 %      97.5 %
ar1  0.4529872330  1.3847478
ar2 -0.5338947632  0.2810423
ma1  0.5525267114  1.4687594
ma2  0.4345131914  1.3694545
ma3  0.3095567052  1.0549870
ma4  0.0009047776  0.4786574
```

ARMA(3,3)  
추가된 모수  
비유의적

- 추가된 모수는 유의적
- 기존의 모수 중 ar2가 비유의적
- ARMA(1,4): 잠정모형



- ARIMA(1,4) 모형 적합

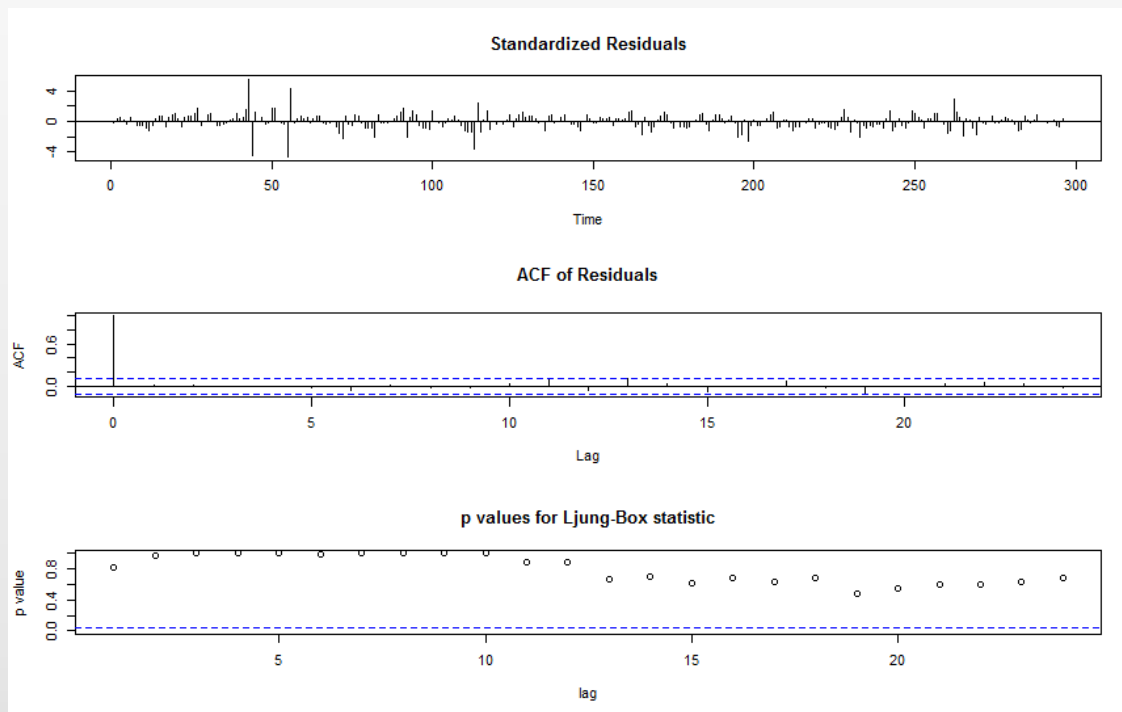
```
> fit2.2 <- arima(gas,order=c(1,0,4),include.mean=FALSE)
> fit2.2

Coefficients:
            ar1         ma1         ma2         ma3         ma4
        0.7769    1.1454    1.0350    0.7846    0.3014
s.e.    0.0442    0.0647    0.0907    0.0862    0.0613

sigma^2 estimated as 0.03352:  log likelihood = 80.01,  aic = -
148.02
```

- ARMA(1,4) 모형의 검진

```
> tsdiag(fit2.2, gof.lag=24)
```



오차항 가정 만족?

- ARMA(1,4) 모형의 과대적합

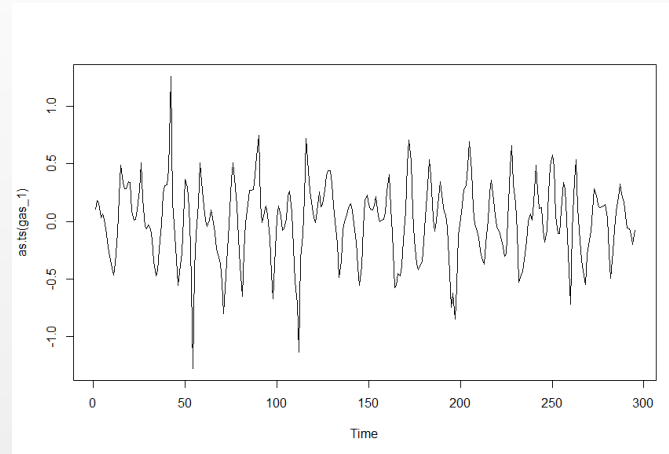
ARMA(1,4) 모형을 최종 모형  
으로 선택 가능

```
> confint(arima(gas,order=c(2,0,4),include.mean=FALSE))
              2.5 %      97.5 %
ar1  0.4529872330  1.3847478
ar2 -0.5338947632  0.2810423
ma1  0.5525267114  1.4687594
ma2  0.4345131914  1.3694545
ma3  0.3095567052  1.0549870
ma4  0.0009047776  0.4786574

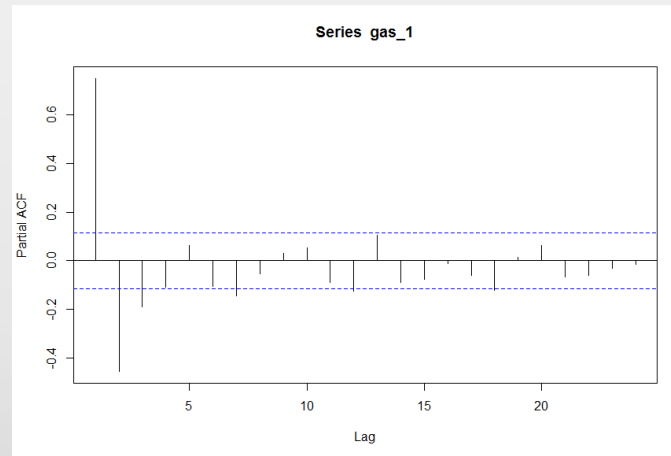
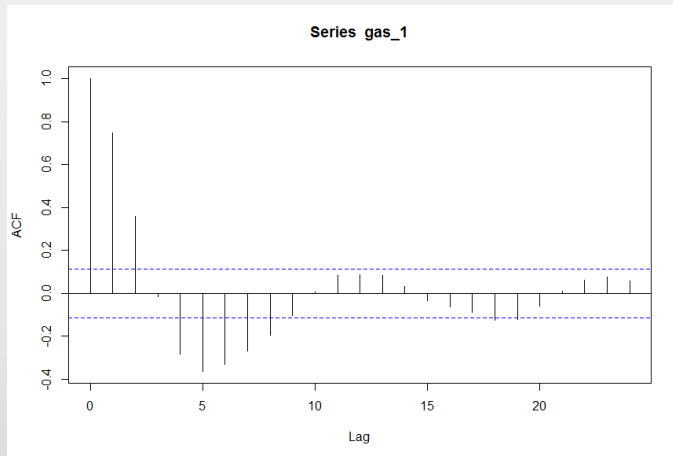
> confint(arima(gas,order=c(1,0,5),include.mean=FALSE))
              2.5 %      97.5 %
ar1  0.6369707  0.8718054
ma1  1.0138346  1.3381123
ma2  0.8326755  1.3535451
ma3  0.5710654  1.1406328
ma4  0.1177378  0.6273530
ma5 -0.1161688  0.2259874
```

## ◆ 차분 차수 $d=1$ 의 경우: 모형 인식

```
> gas_1 <- diff(gas)
> plot(as.ts(gas_1))
> acf(gas_1)
> pacf(gas_1)
```



모형 인식  
ARIMA(3,1,0)



ARIMA(1,1,1)  
ARIMA(2,1,1)  
ARIMA(1,1,2)  
ARIMA(2,1,2)

## 1) ARIMA(3,1,0) 모형 추정 및 검진

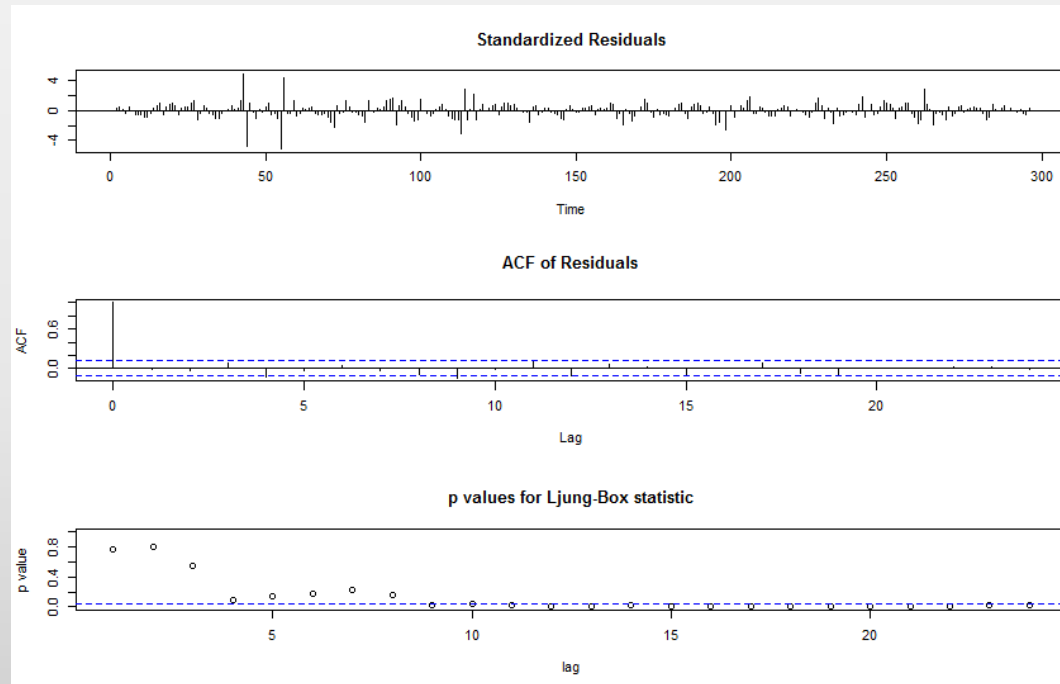
```
> fit3 <- arima(gas,order=c(3,1,0))
> confint(fit3)
```

	2.5 %	97.5 %
ar1	0.8850503	1.10858832
ar2	-0.4028026	-0.08852704
ar3	-0.2996857	-0.07642704

- $d=0$ 에서 절편 비유의적  
→  $d=1$ 에서도 절편 비유의적
- 함수 `arima( )`:  $d \geq 1$  인 경우 절편은 무시됨  
→ 함수 `forecast::Arima( )`

```
> tsdiag(fit3,gof.lag=24)
```

부적합한 모형



## 2) ARIMA(1,1,1), ... , ARIMA(2,1,2) 모형 선택

```
> arima(gas,order=c(1,1,1))$aic  
[1] -93.49603  
> arima(gas,order=c(1,1,2))$aic  
[1] -99.81915  
> arima(gas,order=c(2,1,1))$aic  
[1] -129.7403  
> arima(gas,order=c(2,1,2))$aic  
[1] -128.5871
```

ARIMA(2,1,1) 모형 선택

- ARIMA(2,1,1) 모형의 모수 추정 및 검진

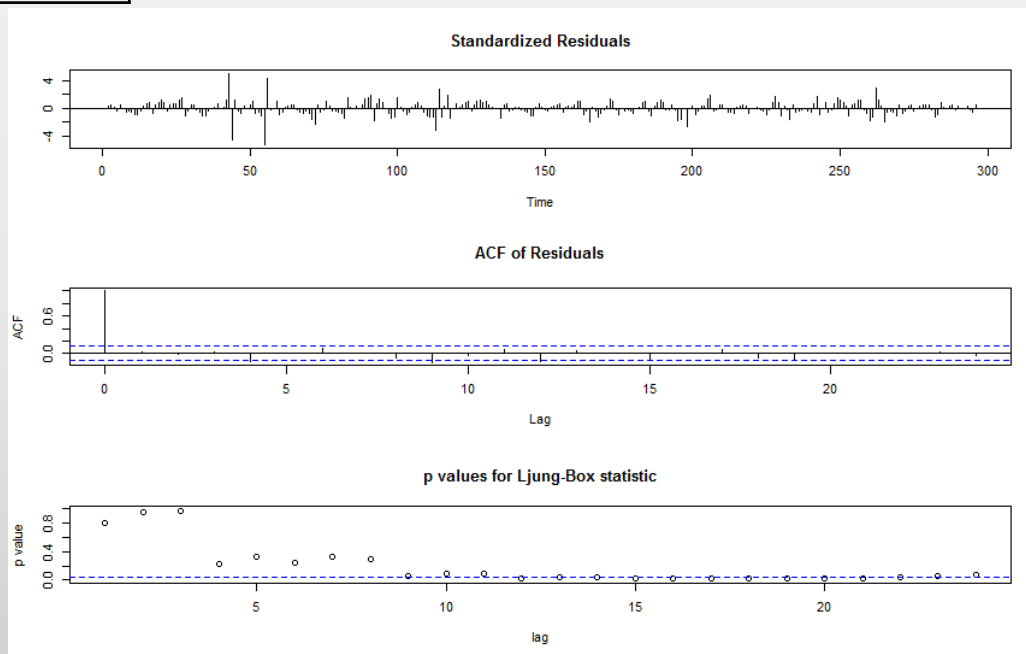
```
> fit4 <- arima(gas, order=c(2,1,1))
> confint(fit4)
```

	2.5 %	97.5 %
ar1	1.2831357	1.6008883
ar2	-0.8300362	-0.6011514
ma1	-0.7067032	-0.2541611

과대적합 모형: 부적절

```
> tsdiag(fit4, gof.lag=24)
```

부적합한 모형



- 참고: 함수 `auto.arima()`로 모형 인식 결과

```
> library(forecast)
> fit5 <- auto.arima(gas)
> fit5
```

Coefficients:

	ar1	ar2	ar3	ma1
	1.9603	-1.3523	0.3310	-0.9861
s.e.	0.0571	0.1017	0.0568	0.0149

sigma^2 estimated as 0.03604: log likelihood=72.21  
AIC=-134.42 AICc=-134.21 BIC=-115.98

```
> confint(fit5)
          2.5 %      97.5 %
ar1  1.8483576  2.0722843
ar2 -1.5515903 -1.1530846
ar3  0.2196371  0.4423471
ma1 -1.0152691 -0.9569575
```

- ARIMA(3,1,1): 검진 단계 통과 못함. 부적합 모형
- 차분 실시 후 적절한 모형 수립 실패
- ARMA(1,4): 최종 모형



- 최종 모형

```
> fit2.2 <- arima(gas,order=c(1,0,4),include.mean=FALSE)
```

```
> fit2.2
```

Coefficients:

	ar1	ma1	ma2	ma3	ma4
	0.7769	1.1454	1.0350	0.7846	0.3014
s.e.	0.0442	0.0647	0.0907	0.0862	0.0613

sigma^2 estimated as 0.03352: log likelihood = 80.01, aic = -148.02

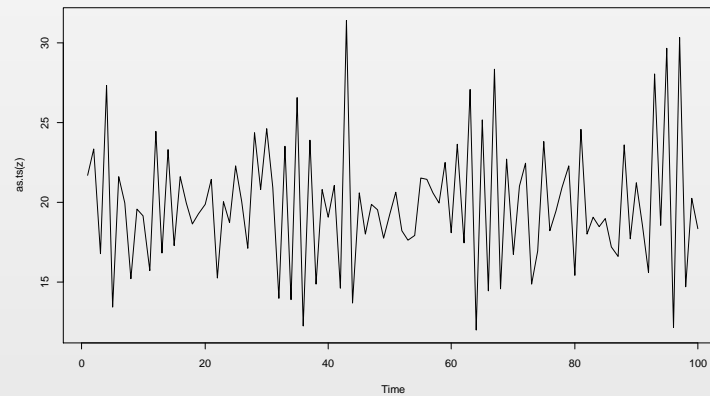
$$Z_t = 0.77Z_{t-1} + \varepsilon_t + 1.14\varepsilon_{t-1} + 1.03\varepsilon_{t-2} + 0.78\varepsilon_{t-3} + 0.30\varepsilon_{t-4}$$

$$\hat{\sigma}^2 = 0.03$$

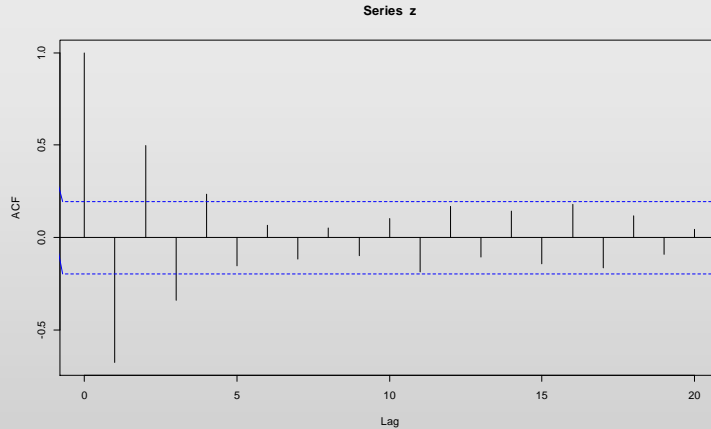
## 예 8-7(312쪽): 데이터 파일 eg8\_7.txt

```
> z<- scan("D:/Data/eg8_7.txt")  
Read 100 items
```

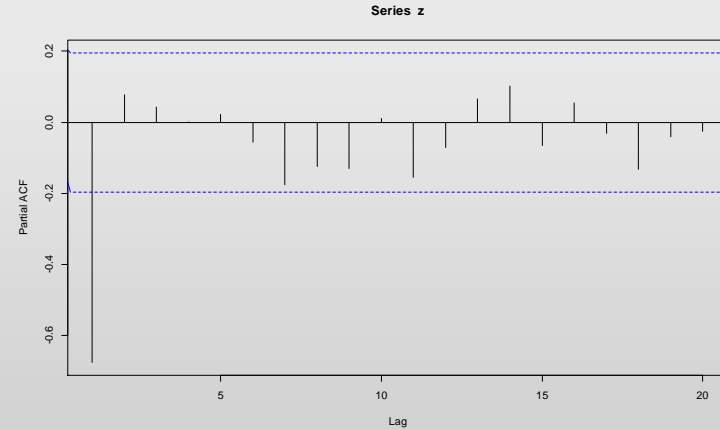
AR(1) 인식



ACF



PACF



- AR(1) 모형 적합 결과

```
> fit <- arima(z, order=c(1,0,0))
> fit

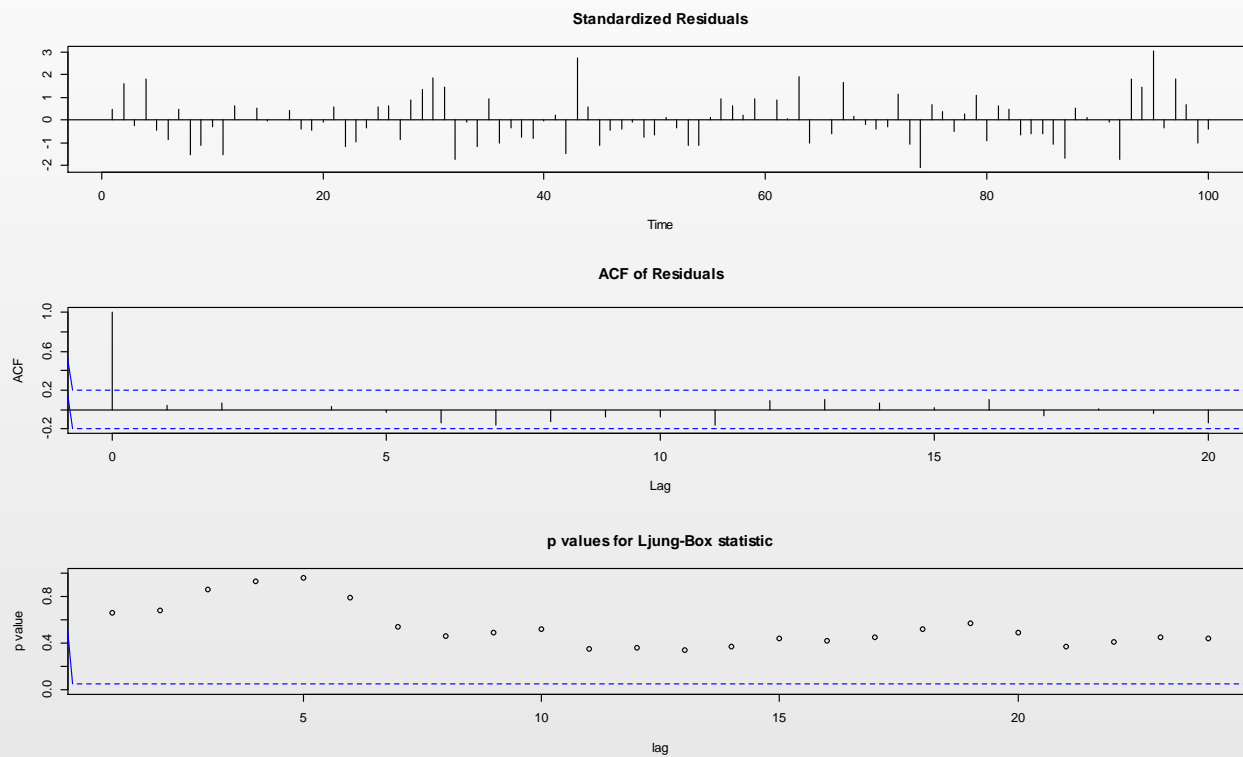
Coefficients:
              ar1      intercept
            -0.6715      19.8312
s.e.          0.0728      0.1776

sigma^2 estimated as 8.744:  log likelihood = -250.61,
aic = 507.22
```

**모형식**

$$Z_t - 19.83 = -0.67(Z_{t-1} - 19.83) + \varepsilon_t$$
$$Z_t = 33.14 - 0.67Z_{t-1} + \varepsilon_t$$
$$\hat{\sigma}^2 = 8.744$$

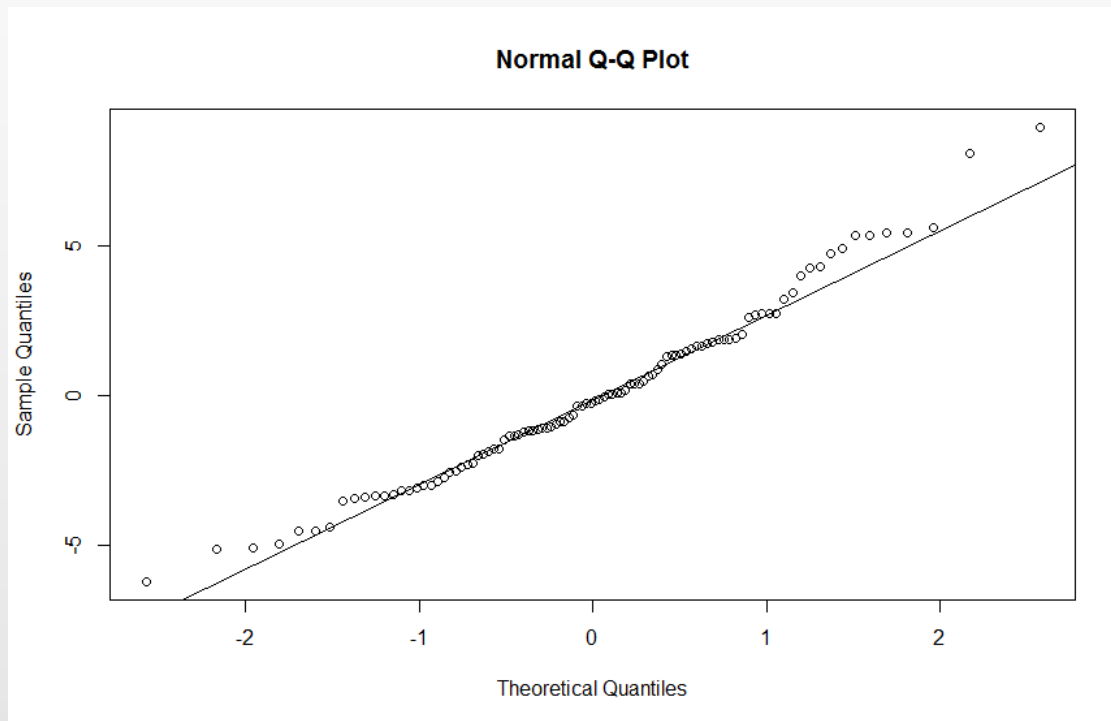
- AR(1) 모형 검진: 잔차분석



문제가 없는 것으로  
보임

- 정규분포 확인

```
> qqnorm(fit$resid); qqline(fit$resid)
```



- AR(1) 모형 검진: 과대적합

```
> confint(arima(z,order=c(2,0,0)))
              2.5 %      97.5 %
ar1      -0.8194184 -0.4268145
ar2      -0.1261271  0.2669024
intercept 19.4595791 20.2066492

> confint(arima(z,order=c(1,0,1)))
              2.5 %      97.5 %
ar1      -0.9126512 -0.5301982
ma1      -0.1807284  0.3661018
intercept 19.4645647 20.2012674
```

두 모형 모두 추가된 모수가 비유의적  
→ AR(1) 모형을 최적모형으로 선택 가능